Report

Student ID：z5192086　　　Name: Pan Luo

Model description：

I use LSTM model to finish this assignment. Because the task is to differentiate positive and negative review on English words. And according to what we have learned, the recurrent neural network is a good way to analyze heavy data set. Once the data and label going through the network, LSTM model can implement different gates to study data in one label. So I use LSTM to build my neural network. And I build three-layer LSTM and connect them together in order to "remember" more long message and avoid long-term dependency. I also apply learning rate decay to avoid over fitting.

Hyperparameter choose:

There are many hyperparameter in LSTM. First, I add drop out layer in every LSTM to avoid over-fitting. I set the keep probability as 0.5. Then I begin to try different parameter which can build my LSTM network better. First, I find my network is always over-fitting on evaluation set. So I try to avoid this. I adjust the number of my LSTM hidden layer, and increase the max words in review because after preprocessing I find there are many long review, if I do not include them, my accuracy on evaluation set will go down.

Final hyperparameter:

batch_size=256,max_words_in_review=300,start_learning_rate=0.01,decay_step= 1000,drop_out_prob=0.5,   lstm_hidden_size=128,